2

OIPE

#2

RAW SEQUENCE LISTING DATE: 08/06/2001 PATENT APPLICATION: US/09/915,873 TIME: 15:34:23

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\08062001\I915873.raw

```
ENTERED
        4 <110> APPLICANT: Stomp, Anne-Marie
                Dickey, Lynn
        6
                Gasdaska, John
        9 <120> TITLE OF INVENTION: Expression of Biologically Active
                Polypeptides in Duckweed
       10
       12 <130> FILE REFERENCE: 40989/237225
C--> 14 <140> CURRENT APPLICATION NUMBER: US/09/915,873
  C--> 14 <141> CURRENT FILING DATE: 2001-07-26
       14 <150> PRIOR APPLICATION NUMBER: US 60/293,330
       15 <151> PRIOR FILING DATE: 2001-05-23
       17 <150> PRIOR APPLICATION NUMBER: US 60/221,705
       18 <151> PRIOR FILING DATE: 2000-07-31
       20 <160> NUMBER OF SEQ ID NOS: 8
       22 <170> SOFTWARE: FastSEQ for Windows Version 4.0
       24 <210> SEQ ID NO: 1
       25 <211> LENGTH: 554
       26 <212> TYPE: DNA
       27 <213> ORGANISM: Zea mays
       29 <400> SEQUENCE: 1
       30 gatcaagtge aaaggteege ettgtttete etetgtetet tgatetgaet aatettggtt 60
       31 tatgattegt tgagtaattt tggggaaage ttegteeaca gttttttttt egatgaacag 120
       32 tgccqcaqtq qcqctqatct tqtatqctat cctqcaatcq tqqtqaactt atqtctttta 180
       33 tateetteae taeeatgaaa agaetagtaa tetttetega tgtaacateg teeageaetg 240
       34 ctattaccgt gtggtccatc cgacagtctg gctgaacaca tcatacgata ttgagcaaag 300
       35 atctatcttc cctgttcttt aatgaaagac gtcattttca tcagtatgat ctaagaatgt 360
       36 tycaacttyc aaggagycyt ttctttcttt gaatttaact aactcyttya ytyycetyt 420
       37 ttctcggacg taaggccttt gctgctccac acatgtccat tcgaatttta ccgtgtttag 480
       38 caagggegaa aagtttgeat ettgatgatt tagettgaet atgegattge ttteetggae 540
       39 ccgtgcagct gcgg
       41 <210> SEQ ID NO: 2
       42 <211> LENGTH: 498
       43 <212> TYPE: DNA
       44 <213> ORGANISM: Artificial Sequence /
       46 <220> FEATURE:
       47 <223> OTHER INFORMATION: Duckweed codon optimized nucleotide sequence
                encoding human alpha-2B interferon >
       50 <221> NAME/KEY: CDS
       51 <222> LOCATION: (1)...(498)
       53 <400> SEQUENCE: 2
                                                                             48
       54 tgc gae etc ecc cag acc cac age etc ggg tec egc egc acc etc atg
       55 Cys Asp Leu Pro Gln Thr His Ser Leu Gly Ser Arg Arg Thr Leu Met
                                               10
                                                                             96
       58 ctg ctg gcg cag atg cgc cgc atc tcg ctc ttc agc tgc ctg aag gac
       59 Leu Leu Ala Gln Met Arg Arg Ile Ser Leu Phe Ser Cys Leu Lys Asp
                       20
                                           25
       62 ege cae gae tte gge tte eeg eag gag gag tte gge aae eag tte eag
                                                                             144
       63 Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Gly Asn Gln Phe Gln
```

RAW SEQUENCE LISTING DATE: 08/06/2001 PATENT APPLICATION: US/09/915,873 TIME: 15:34:23

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\08062001\I915873.raw

			35					40					4.5				
64 66	aag	acc		acq	atc	ccc	ata		cac	σασ	atσ	atc		caq	atc	ttc	192
	Lys																
68	_1_	50					55					60					
	aac		t.t.c	aσc	acc	aaσ		aσc	t.ca	acc	acc		gac	σασ	acc	at.a	240
	Asn																
72	65					70					75					80	
. –	ctc	gac	aad	ttc	tac		αaα	cta	tac	caσ		ctc	aac	gac	cta		288
	Leu																
76			-1-		85				- 1 -	90				1	95		
	gcg	tac	ata	atc	cag	aaa	att	aaa	att		gag	acq	cca	cta	atq	aaq	336
	Ala																
80		-1-		100		1		1	105					110		-1-	
	gag	gac	agc		ctc	acc	ata	cac		tac	ttc	caq	cac		acq	ctc	384
	Glu																
84	014	TICF	115					120	-10	-1-		0	125				
	tac	ct.c		σασ	aaσ	aaσ	tac		cca	tac	acc	taa		atc	att	cac	432
	Tyr		_		_	_		_	-	_	-			_	_	-	
88	-1-	130	-1-		-10	-10	135			010		140				9	
	gcc		atc	at.a	cac	t.cc		agc	cta	age	acc		ctc	caσ	σασ	agc	480
	Ala							-	-	-				-		-	
	145				5	150					155			0 _ 11		160	
	ctc	cac	t.cc	aaσ	σασ												498
	Leu	_		_		*											.,,
96		,		-1-	165												
99	<210	)> SE	O TI	NO.	3												
			, TI	, 110.													
100	<21																
	<21	.1> I	ENG	H: 9	96												
102	<21 <21	1> I .2> T	ENGT	TH: S	)6 1	za s	ativ	ra									
102 102	<21	.1> I .2> T .3> C	ENGT YPE: RGAN	TH: 9 DNA	96 A : Ory	za s	ativ	ra									
102 102 104	<pre>21   &lt;21   &lt;21   &lt;21   &lt;40 </pre>	1> I .2> T .3> C	ENGT YPE: RGAN EQUI	TH: S DNA NISM: ENCE:	96 Ory 3				ıc aa	ıqcad	ettec	c tct	ccct	igto	cqto	ctcat	cc 60
102 102 104 105	<pre></pre>	1> I 2> 7 3> 0 00> s	ENGT YPE: RGAN EQUI	TH: S DNA NISM: ENCE: tcct	06 Cory 3 Cgaac	eac g	ratgg	rtcaa				e tet	ccct	igtc	cgto	cctcat	cc 60 96
102 102 103 105	<pre>21   &lt;21   &lt;21   &lt;21   &lt;40 </pre>	1> I 2> I 3> C 00> S catgo	ENGT TYPE: ORGAN SEQUE Cagg	TH: 9 INF INCE: tcct ggct	06 Ory 3 Igaad Igagd	eac g	ratgg	rtcaa				c tct	ceet	igtc	cgto	ectcat	
102 102 103 103 106	<pre></pre>	1> I 2> I 3> 0 00> S catgo	ENGT TYPE: RGAN EQUE agg tcg	TH: S ISM: NISM: ENCE: tcct ggct D NO	Ory 3 cgaac cgagc 0: 4	eac g	ratgg	rtcaa				c tct	ccct	igtc	cgto	ectcat	
102 102 103 105 106 108	<pre>21 &lt;21 &lt;21 &lt;21 &lt;21 &lt;40 according to 3 &lt;21</pre>	1> I 2> T 3> C 00> S atgo cctco .0> S	ENGT TYPE: DRGAN EEQUE Eagg Etcg EEQ I	TH: 9 DNA NISM: ENCE: tcct ggct D NC	Off Ory 3 cgaac cgago : 4	eac g	ratgg	rtcaa				: tct	ccct	igto	cgtc	ectcat	
102 102 103 106 108 108	<pre>21 &lt;21 2 &lt;21 4 &lt;40 5 acc 6 gtc 8 &lt;21 9 &lt;21</pre>	1> I 2> T 3> C 00> S catgo cctco .0> S .1> I .2> T	ENGT TYPE: DRGAN SEQUE SEQUE SEQUE SEQUE SEQUENCT	TH: 9 DNA NISM: ENCE: tcct ggct D NC TH: 1	06 Cory 3 Cgaac Cgagc 0: 4	eac g eag c	ratgo aaco	rtcaa				c tct	ccct	igtc	cgto	ectcat	
103 103 103 103 106 108 110	<pre>3</pre>	1> I 2> T 3> C 00> S catgo cctco 0> S 1> I .2> T	ENGT TYPE: PRGAN EQUI- Cagg ELCG ELCG TYPE: PRGAN	TH: 9 DNA NISM: ENCE: tect gget D NO TH: 1 PRI	Off Ory 3 Cgaac Cgagc O: 4 88	eac g eag c	ratgo aaco	rtcaa				tct	cecet	igtc	cgto	cctcat	
102 103 103 106 108 109 110	21 <21 <21 <21 <21 <21 <21 <21 <40 <40 <5 acc <5 gtc <3 <21 <21 <21 <21 <21 <21 <40 <40 <40 <40 <40 <40 <40 <40 <40 <40	1> I 2> T 3> C 00> S catgo cctco .0> S .1> I .2> T	ENGT TYPE: DRGAN EQUI EAGG ECQ I ENGT TYPE: DRGAN EQUI	TH: 9 DNA NISM: ENCE: tcct ggct D NC TH: 1 PRI NISM: UNCE:	06 Cgaac cgagc ): 4 .88 Hom	eac g eag d	ratgo aaco	rtcaa etcac	ec go	cggc	2						
102 103 103 106 108 109 110	21 <21 <21 <21 <21 <21 <21 <21 <21 <21 <	1> I 2> T 3> C 00> S catgo cctco .0> S .1> I .2> T	ENGT TYPE: DRGAN EQUI EAGG ECQ I ENGT TYPE: DRGAN EQUI	TH: 9 DNA NISM: ENCE: tcct ggct D NC TH: 1 PRI NISM: UNCE:	06 Cgaac cgagc ): 4 .88 Hom	eac g eag d	ratgo aaco	rtcaa etcac	ec go	cggc	2					cctcat Cys	
101 102 103 106 108 109 110 113 114 115	0 <21	1> I 2> T 3> C 00> S catgo cotco 0> S 1> I .2> T .3> C 00> S	LENGT TYPE: ORGAN EQUI- EAGG EQUI- DRGAN EQUI- EQUI- LENGT DRGAN	TH: 9 DNA NISM: ENCE: tcct ggct D NC TH: 1 PRI NISM: ENCE: Thi	Ory 3 cgaac cgagc 1: 4 88 Hom 4 Phe	eac g eag c no sa	atgg aacc	rtcaa etcac us	ec go	ecggo Ala	Lev	ı Lev	ı Val	Leu	ı Ser 15	· Cys	
101 102 103 106 108 109 110 113 114 115	2 < 21	1> I 2> T 3> C 00> S catgo cotco 0> S 1> I .2> T .3> C 00> S	LENGT TYPE: ORGAN EQUI- EAGG EQUI- DRGAN EQUI- EQUI- LENGT DRGAN	TH: 9 DNA NISM: ENCE: tcct ggct D NC TH: 1 PRI NISM: ENCE: Thi	Ory 3 cgaac cgagc 1: 4 88 Hom 4 Phe	eac g eag c no sa	atgg aacc	rtcaa etcac us	ec go	ecggo Ala	Lev	ı Lev	ı Val	Leu	ı Ser 15		
102 102 103 106 108 108 113 114 115 116	2 <21	1> I 2> T 3> C 00> S catgo cctco .0> S .1> I .2> T .3> C .0> S	LENGTYPE: PRGANGEQUE CAGGE CAG	TH: 9 F DNA  NISM: ENCE: tcct ggct D NC  CH: 1 PRI  NISM: ENCE: Thi  Cys 20	Official Control of the Control of t	eac g cag d no sa e Ala	atgg aacc pier Leu	rtcaa etcac is Leu Cys	val Asp 25	Ala 10 Leu	Leu Pro	ı Leu o Glr	ı Val	Leu His	s Ser 15 Ser	· Cys	
102 102 103 106 108 108 113 114 115 116	0 <21 <21 <21 <40 6 acc 6 gtc 8 <21 0 <21 <21 <21 5 <40 6 Lys	1> I 2> T 3> C 00> S catgo cctco .0> S .1> I .2> T .3> C .0> S	LENGTYPE: PRGANGEQUE CAGGE CAG	TH: 9 F DNA  NISM: ENCE: tcct ggct D NC  CH: 1 PRI  NISM: ENCE: Thi  Cys 20	Official Control of the Control of t	eac g cag d no sa e Ala	atgg aacc pier Leu	rtcaa etcac is Leu Cys	val Asp 25	Ala 10 Leu	Leu Pro	ı Leu o Glr	ı Val	Leu His	s Ser 15 Ser	Cys Leu	
102 103 104 105 106 108 110 111 113 114 115 116 117	2 < 21	1> II 2> II 3> C 00> Seatgone (Control of the Control of the Co	ENGTYPE: PRGANTEQUE EAGGE EAGG	TH: 9 FINE FINE FINE FINE FINE FINE FINE FINE	Office Ory 3 cgaaccgagc 2: 4 cm 4 cm Phe 5 cm Seria Thr	eac gag como sa e Ala	pier Leu Gly	tcaa tcaa s Leu Cys Leu 40	Val Asp 25 Leu	Ala 10 Leu	Leu Pro	Leu Glr Met	Val Thr Arg 45	Leu His 30 Arg	Ser 15 Ser J Ile	Cys Leu	
102 103 104 105 106 108 110 111 113 114 115 116 117	2 < 21	1> II 2> II 3> C 00> Seatgone (Control of the Control of the Co	ENGTYPE: PRGANTEQUE EAGGE EAGG	TH: 9 FINE FINE FINE FINE FINE FINE FINE FINE	Office Ory 3 cgaaccgagc 2: 4 cm 4 cm Phe 5 cm Seria Thr	eac gag como sa e Ala	pier Leu Gly	tcaa tcaa s Leu Cys Leu 40	Val Asp 25 Leu	Ala 10 Leu	Leu Pro	Leu Glr Met	Val Thr Arg 45	Leu His 30 Arg	Ser 15 Ser J Ile	Cys Leu	
103 104 105 106 108 109 110 111 113 114 115 116 117 118 120 121	2 < 21	1> II 2> II 3> C 00> Seatgon ceteco 0> Seatgon 1> II 3> C 10> Seatgon 1> II 3> C 1> II 3> C 1> Seatgon 1> II 3> C 1> II 3> C 1> Seatgon 1> II 3> C 1> Seatgon 1- Sea	ENGTYPE: DRGANGEQUE CAGGETOGE ENGTYPE: DRGANGEQUE CAGGETOGE CAGGETOGE CAGGETOGETOGETOGETOGETOGETOGETOGETOGETOGET	TH: 9 DNA NISM: ENCE: toot gget D NO TH: 1 PRI NISM: CNCE: Thi Cys 20 Arg	Official Control Contr	eac gag como sa e Alar Val	pier Gly Met	tcaa tcaa Leu Cys Leu 40	Val Asp 25 Leu	Ala 10 Leu Ala Asp	Leu Pro Gln Phe	Leu O Glr Met Gly 60	Val Thr Arg 45	Leu 30 Arg	Ser 15 Ser Ile	Cys Leu	
103 104 105 106 108 109 110 113 114 115 116 117 118 120 121 122	2 < 21	1> II 2> II 3> C 00> Seatgon ceteco 0> Seatgon 1> II 3> C 10> Seatgon 1> II 3> C 1> II 3> C 1> Seatgon 1> II 3> C 1> II 3> C 1> Seatgon 1> II 3> C 1> Seatgon 1- Sea	ENGTYPE: DRGANGEQUE CAGGETOGE ENGTYPE: DRGANGEQUE CAGGETOGE CAGGETOGE CAGGETOGETOGETOGETOGETOGETOGETOGETOGETOGET	TH: 9 DNA NISM: ENCE: toot gget D NO TH: 1 PRI NISM: CNCE: Thi Cys 20 Arg	Official Control Contr	eac gag como sa e Alar Val	pier Gly Met	tcaa tcaa Leu Cys Leu 40	Val Asp 25 Leu	Ala 10 Leu Ala Asp	Leu Pro Gln Phe	Leu O Glr Met Gly 60	Val Thr Arg 45	Leu 30 Arg	Ser 15 Ser Ile	Cys Leu Ser	
101 102 103 106 108 109 110 111 113 114 115 116 117 120 121 122 123	0 <21	1> I 12> I 10> Seatgon 10> S	ENGTYPE: DRGANGEQUE CAGE CAGE CAGE CAGE CAGE CAGE CAGE CAG	TH: 9 DNA NISM: ENCE: toot gget D NO PRO IISM: ENCE: Thi Cys 20 J Arg	Office Ory 3 cgaaccgagco: 4 .888	cac gad constant search	pier Gly Met	tcaa stcaa Leu Cys Leu 40 Arg	Val Asp 25 Leu His	Ala 10 Leu Ala Asp	Leu Pro Glm Phe Thr 75	Leu O Glr Met Gly 60	Val Thr Arg 45 Phe	Leu His 30 Arg Pro	Ser 15 Ser Ile Glr	Cys Leu Ser Glu His	
101 102 103 106 108 109 110 111 113 114 115 116 117 120 121 122 123	0 <21	1> I 12> I 10> Seatgon 10> S	ENGTYPE: DRGANGEQUE CAGE CAGE CAGE CAGE CAGE CAGE CAGE CAG	TH: 9 DNA NISM: ENCE: toot gget D NO PRO IISM: ENCE: Thi Cys 20 J Arg	Office Ory 3 cgaaccgagco: 4 .888	cac gad constant search	pier Gly Met	tcaa stcaa Leu Cys Leu 40 Arg	Val Asp 25 Leu His	Ala 10 Leu Ala Asp	Leu Pro Glm Phe Thr 75	Leu O Glr Met Gly 60	Val Thr Arg 45 Phe	Leu His 30 Arg Pro	Ser 15 Ser Ile Glr	Cys Leu Ser Glu His	

RAW SEQUENCE LISTING DATE: 08/06/2001 PATENT APPLICATION: US/09/915,873 TIME: 15:34:23

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\08062001\1915873.raw

126 Ala Ala Trp Asp Glu Thr Leu Leu Asp Lys Phe Tyr Thr Glu Leu Tyr 100 105 128 Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Ile Gln Gly Val Gly Val 120 130 Thr Glu Thr Pro Leu Met Lys Glu Asp Ser Ile Leu Ala Val Arg Lys 131 130 135 132 Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Lys Glu Lys Lys Tyr Ser Pro 150 155 134 Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser Leu 170 165 136 Ser Thr Asn Leu Gln Glu Ser Leu Arg Ser Lys Glu 137 180 185 140 <210> SEQ ID NO: 5 141 <211> LENGTH: 165 142 <212> TYPE: PRT 143 <213> ORGANISM: Homo sapiens 145 <400> SEQUENCE: 5 146 Cys Asp Leu Pro Gln Thr His Ser Leu Gly Ser Arg Arg Thr Leu Met 5 10 148 Leu Leu Ala Gln Met Arg Arg Ile Ser Leu Phe Ser Cys Leu Lys Asp 20 25 150 Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Gly Asn Gln Phe Gln 35 40 152 Lys Ala Glu Thr Ile Pro Val Leu His Glu Met Ile Gln Gln Ile Phe 55 154 Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu Thr Leu 155 65 70 156 Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu Glu 157 85 158 Ala Cys Val Ile Gln Gly Val Gly Val Thr Glu Thr Pro Leu Met Lys 100 105 160 Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr Leu 120 162 Tyr Leu Lys Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val Arg 130 135 140 164 Ala Glu Ile Met Arg Ser Phe Ser Leu Ser Thr Asn Leu Gln Glu Ser 165 145 150 155 166 Leu Arg Ser Lys Glu 167 170 <210> SEQ ID NO: 6 171 <211> LENGTH: 31 172 <212> TYPE: PRT 173 <213> ORGANISM: Oryza sativa 175 <400> SEQUENCE: 6 176 Met Gln Val Leu Asn Thr Met Val Asn Lys His Phe Leu Ser Leu Ser 177 1 5 10 178 Val Leu Ile Val Leu Leu Gly Leu Ser Ser Asn Leu Thr Ala Gly 25 182 <210> SEO ID NO: 7

RAW SEQUENCE LISTING DATE: 08/06/2001 PATENT APPLICATION: US/09/915,873 TIME: 15:34:23

Input Set : A:\Seqlist.txt

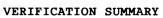
Output Set: N:\CRF3\08062001\I915873.raw

183 <211> LENGTH: 31 184 <212> TYPE: PRT 185 <213> ORGANISM: Artificial Sequence 187 <220> FEATURE: 188 <223> OTHER INFORMATION: Modified rice alpha-amylase signal peptide 190 <400> SEQUENCE: 7 191 Met Gln Val Leu Asn Thr Met Val Asn Lys His Phe Leu Ser Leu Ser 5 193 Val Leu Ile Val Leu Thr Val Leu Ser Ser Asn Leu Thr Ala Gly 194 20 25 197 <210> SEQ ID NO: 8 198 <211> LENGTH: 21 199 <212> TYPE: PRT 200 <213> ORGANISM: Arabidopsis thaliana 202 <400> SEQUENCE: 8 203 Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Ser 204 1 5

205 Leu Ser Ser Ala Glu

20

206



DATE: 08/06/2001 PATENT APPLICATION: US/09/915,873 TIME: 15:34:24

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\08062001\I915873.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application No

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date